Revue de statistique appliquée

M. Roche

Fiabilité d'un matériel

Revue de statistique appliquée, tome 10, nº 3 (1962), p. 29-58

<http://www.numdam.org/item?id=RSA_1962__10_3_29_0>

FIABILITÉ D'UN MATÉRIEL

Mlle M. ROCHE

Ingénieur à la Société d'Économie et de Mathématiques Appliquées

La fiabilité (ou sûreté de fonctionnement) d'un matériel se définit comme la probabilité pour celui-ci de fonctionner dans des conditions d'utilisation données et pendant un temps donné.

A cette notion de fiabilité se trouve étroitement associée celle d'avarie. En effet, du mode d'apparition des avaries auquel est soumis le matériel, dépend sa fonction fiabilité.

Nous étudierons donc les diverses formes d'avaries possibles pour en déduire la fonction fiabilité correspondante.

Les avaries d'un matériel sont dues à un ou plusieurs ensembles superposés de causes plus ou moins complexes. On distingue en général trois grandes causes d'avaries :

- la forte variation accidentelle d'une ou plusieurs conditions d'utilisation (panne d'un ordinateur due à une forte élévation de température).

- la fatigue consécutive à un nombre répété de variations des conditions d'utilisation (vibrations...);

- l'usure d'un élément au cours du temps (jeu..).

Ces causes peuvent d'ailleurs agir simultanément.

Nous allons passer en revue les modèles les plus souvent utilisés pour représenter ces différents types d'avaries.

I - DIFFERENTES FORMES DE DISTRIBUTIONS DE LA DUREE DE VIE D'UN MATERIEL. FONCTIONS DE FIABILITE -

Un paramètre fondamental de ces distributions est le taux d'avarie $g(t)$, [$g(t)\,dt$ représentant la probabilité d'avarie dans l'intervalle $(t, t + dt)$, sachant qu'il est encore en vie à l'instant t],

d'expression :

$$g(t) = \frac{f(t)}{1 - F(t)}$$

où $F(t)$ représente la fonction de répartition de la durée de vie et $f(t)$ sa densité de probabilité.

A - Supposons que le matériel soit soumis à un seul ensemble de conditions extérieures et que celles-ci puissent avoir des pointes (c'est-à-dire de fortes variations).

Supposons d'autre part que l'apparition de ces pointes se fasse suivant une loi poissonnienne de taux λ.

a) Si à chaque pointe correspond une avarie, la fonction fiabilité d'un tel matériel, définie comme la probabilité pour la variable aléatoire durée de vie, T, d'être supérieure à une quantité t, est donnée par :

$$R(t) = Pr\{T > t\} = Pr\{0 \text{ pointe}\} = e^{-\lambda t}, \qquad t \geqslant 0 \tag{1}$$

La probabilité d'avoir une avarie dans l'intervalle (t, t + dt) est donc :

$$f(t)\,dt = \lambda e^{-\lambda t} \qquad t \geqslant 0 \tag{1'}$$

(λ, taux d'avarie constant est égal à l'inverse de la durée de vie moyenne $\vartheta = \int_0^\infty t\, f(t)\, dt$.

b) Si k pointes sont nécessaires pour entraîner l'avarie (phénomène de fatigue) la fonction fiabilité devient :

$$R_k(t) = Pr\{T > t\} = \text{Probabilité d'avoir moins de k pointes dans } (o, t)$$

$$= \sum_{j=o}^{k=1} e^{-\lambda t} \frac{(\lambda t)^j}{j!} \tag{2}$$

et la probabilité d'avoir une avarie dans l'intervalle (t, t + dt) s'écrit :

$$f_k(t) = Pr\{(k-1) \text{ pointes dans } (o, t) \text{ et } 1 \text{ pointe dans } (t, t + dt)\}$$

$$f_k(t) = \frac{\lambda^k t^{k-1}}{(k-1)!} e^{-\lambda t} \qquad t \geqslant 0 \tag{2'}$$

c) On peut envisager une généralisation de b) en faisant k variable. Autrement dit, l'apparition d'une pointe entraîne une avarie avec une certaine probabilité p.

Le taux d'apparition des avaries est alors λp et la fonction fiabilité devient :

$$R(t) = e^{-\lambda p t} \qquad t \geqslant 0 \tag{3}$$

De même la probabilité d'avoir une avarie dans l'intervalle (t, t + dt) a pour expression :

$$f(t)\,dt = \lambda p\, e^{-\lambda p t}\,dt \quad , \quad t \geqslant 0 \tag{3'}$$

Ce qui précède se généralise facilement au cas où le matériel est soumis à n ensembles E_i de conditions extérieures caractérisés chacun par un taux d'apparition des pointes λ_i et par une probabilité conditionnelle d'avarie sous l'effet d'une pointe p_i.

- Si les n ensembles interviennent simultanément :

$$R(t) = \prod_{i=1}^{n} e^{-\lambda_i p_i t} = e^{-\sum_{i=1}^{n} \lambda_i p_i t} \qquad t \geqslant 0 \tag{4}$$

et :

$$f(t) = \left(\sum_{i=1}^{n} \lambda_i p_i\right) e^{-\sum_{i=1}^{n} \lambda_i p_i t} \qquad t \geqslant 0 \tag{4'}$$

- Si l'ensemble E_i a une probabilité c_i d'intervenir :

$$R(t) = \sum_{i=1}^{n} c_i \; e^{-\lambda_i p_i t} \qquad t \geqslant 0 \tag{5}$$

$$f(t) = \sum_{i=1}^{n} c_i \lambda_i p_i e^{-\lambda_i p_i t} \qquad t \geqslant 0 \tag{5'}$$

d) Considérons maintenant le cas où la probabilité conditionnelle d'avarie sous l'effet d'une pointe (notée p précédemment) dépend de la variable temps (phénomènes d'usure ou de fatigue).

Soit p(t) cette probabilité.

Le taux d'apparition des avaries n'est alors plus constant, puisque égal à $\lambda p(t)$ à l'instant t.

Posons $g(t) = \lambda p(t)$

g(t), par définition même, peut se mettre sous la forme :

$$g(t) = \frac{f(t)}{R(t)}$$

ce qui s'écrit encore :

$$- g(t) = d\,[\log R(t)]$$

or :

$$R(o) = 1$$

donc :

$$R(t) = e^{-\int_0^t g(\tau)\, d\tau} \qquad t \geqslant 0$$

et :

$$f(t) = g(t) \; e^{-\int_0^t g(\tau)\, d\tau} \qquad t \geqslant 0$$

Dans le cas considéré plus haut, g(t) est égal à $\lambda p(t)$ donc :

$$R(t) = e^{-\lambda \int_0^t p(\tau)\, d\tau} \qquad t \geqslant 0 \tag{6}$$

et :

$$f(t) = \lambda p(t) \; e^{-\lambda \int_0^t p(\tau)\, d\tau} \qquad t \geqslant 0 \tag{6'}$$

La loi de Weibull utilisée assez souvent pour représenter des phénomènes d'usure correspond au cas particulier : $g(t) = k\,\lambda^k t^{k-1}$ où λ et k sont deux paramètres à déterminer.

Pour $g(t) = \alpha\beta e^{\beta t}$ on retrouve la loi intervenant dans la théorie des valeurs extrêmes.

Montrons le à l'aide d'un processus d'usure particulier : la corrosion.

Supposons que la corrosion se traduise par la production de creux dans la matière étudiée. Le matériel sera dit avarié si le creux se transforme en trou. Ce trou est supposé se produire à l'endroit où la matière est la plus profondément attaquée. Si t_i désigne le temps d'apparition d'un trou à l'endroit du $i^{\text{ème}}$ creux, le temps d'apparition d'une avarie est $\min_i \{t_i\}$.

Si l'on admet que la variable aléatoire : profondeur du creux suit une loi exponentielle et que le temps d'apparition d'un trou pour un creux donné est linéairement dépendant de la différence entre l'épaisseur de la matière attaquée et la profondeur de ce creux, la théorie des valeurs extrêmes conduit à une distribution de la durée de vie du matériel de la forme : $\alpha e^t e^{-\alpha e^t}$ $(\beta = 1)$.

Enfin on peut rencontrer des courbes d'avarie avec un taux linéairement dépendant du temps.

C'est le cas en aéronautique par exemple (travaux de l'Ingénieur en Chef R. Descamps) où pour certains types de matériel, les avaries peuvent être attribuées à deux systèmes superposés de causes :

- l'un (causes externes), phénomène de fatigue (vibrations...) conduisant à des avaries strictement aléatoires et sans relation avec l'âge des matériels, d'où un taux d'avarie constant : r_o,

- l'autre résultant de processus d'usure (jeu...) et se traduisant par un taux d'avarie croissant avec l'âge des matériels. Si on admet que cette croissance est linéaire, la variation du taux d'avarie avec l'âge t d'un matériel de type donné est de la forme :

$$g(t) = r_o + r_1^2 t$$

(r_1^2 est introduit par raison d'homogénéité).

On a donc :

$$R(t) = e^{-\int_o^t (r_o + r_1^2 \tau)\, d\tau} = e^{-r_o t - r_1^2 \frac{t^2}{2}} \qquad t \geqslant 0$$

et :

$$f(t) = (r_o + r_1^2 t)\, e^{-r_o t - r_1^2 \frac{t^2}{2}} \qquad t \geqslant 0$$

e) Reprenons le cas où l'avarie ne peut se produire (si elle se produit) que lors de la $k^{\text{ième}}$ pointe consécutive.

Soit p(t) la probabilité avec laquelle elle se produit pour une suite donnée de k pointes.

Les formules (2) et (2') se généralisent et deviennent :

$$R_k(t) = \sum_{j=o}^{k-1} e^{-\lambda P(t)} \frac{[\lambda P(t)]^j}{j!} \qquad t \geqslant 0 \tag{7}$$

$$f_k(t) = \frac{[\lambda P(t)]^{k-1}}{(k-1)!} \lambda p(t) e^{-\lambda P(t)} \qquad t \geqslant 0 \tag{7'}$$

où $P(t) = \int_0^t p(\tau)\, d\tau$.

B - Une généralisation possible du processus de Poisson est le processus dit "de naissance et de mort".

Weiss a montré que certains types d'avaries mécaniques comme celles se produisant pour des filaments polymères orientés soumis à l'action de forces de tension, pouvaient être représentés par un processus de "mort pure" ("pure death").

Plus précisément, considérons N_0 fibres, chacune d'entre elles étant indépendante des $(N_0 - 1)$ autres, susceptibles de défaillance sous l'action de charges. On aura une avarie lorsque les N_0 fibres seront défaillantes.

Soit $\phi(t)\, \Delta t$ la probabilité pour qu'une fibre ait une défaillance dans l'intervalle $(t, t + \Delta t)$.

On a alors :

$$R(t) = 1 - \Pr\{\text{zéro fibre non avariée en } t\}$$

$$R(t) = 1 - \left\{1 - \exp.\left[-\int_0^t \phi(\tau)\, d\tau\right]\right\}^{N_0} \qquad t \geqslant 0 \tag{8}$$

et :

$$f(t) = N_0 \phi(t) \exp\left[-\int_0^t \phi(\tau)\, d\tau\right] \left\{1 - \exp\left[-\int_0^t \phi(\tau)\, d\tau\right]\right\}^{N_0 - 1} \tag{8'}$$

Supposons maintenant que la probabilité de défaillance d'une fibre dépende de l'état des autres fibres (les charges auxquelles sont soumises les N_0 fibres étant constantes quel que soit l'état de celles-ci, les fibres non défaillantes à l'instant t auront une probabilité de défaillance d'autant plus grande qu'elles seront moins nombreuses).

Plus précisément, soit $\frac{N_0}{N_0 - n} \phi(t)\, \Delta t$ la probabilité pour qu'une fibre ait une défaillance dans l'intervalle de temps $(t, t + \Delta t)$ sachant que $n (\leqslant N_0)$ fibres sont déjà défaillantes.

On a alors :

$$R(t) = \sum_{j=0}^{N_0 - 1} \frac{\left[\int_0^t N\ \phi(\tau)\, d\tau\right]^j}{j!} \exp\left[-\int_0^t N_0\, \phi(\tau)\, d\tau\right] \qquad t \geqslant 0 \tag{9}$$

et :

$$f(t) = \left[\int_0^t N_0\, \phi(\tau)\, d\tau\right]^{N_0 - 1} \frac{N_0\, \phi(t)}{(N_0 - 1)!} \left[\exp - \int_0^t N_0\, \phi(\tau)\, d\tau\right] \qquad t \geqslant 0 \tag{9'}$$

On remarquera la similitude de (9) et (9') avec respectivement (7) et (7').

Il est intéressant de noter également que Birnbaum et Saunders arrivent à la même loi pour décrire la distribution des durées de vie de structures à multiples composantes (m) soumises à l'action de charges dynamiques. Ils partent de l'hypothèse qu'au cours du temps, l'une après l'autre les composantes sont défaillantes et qu'il existe un nombre critique $k(< m)$ de défaillances tel que l'ensemble est avarié lorsque ce nombre est atteint.

La probabilité conditionnelle d'avarie quand j composantes sont déjà défaillantes à l'instant t est prise de la forme : $\frac{p(t)}{m - j}$. On obtient alors une densité de probabilité de forme analogue à celle trouvée précédemment.

II - TEST DE LA FORME EXPONENTIELLE DE LA DISTRIBUTION DES DUREES DE VIE D'UN MATERIEL -

Nous venons de voir diverses distributions de la variable aléatoire : durée de vie d'un matériel, correspondant à différents modes d'avaries.

La plus couramment étudiée dans la littérature américaine (notamment par B. Epstein) est la loi exponentielle d'expression :

$$f(t) = \lambda e^{-\lambda t}, \qquad t \geqslant 0 \qquad (1)$$

à taux d'avarie constant.

Sa simplicité donne lieu en effet à des calculs relativement abordables.

De plus, elle décrit de façon satisfaisante les distributions des durées de vie de la plupart des matériels électriques et électroniques pour lesquels le phénomène d'usure peut être considéré comme pratiquement inexistant.

Elle apparaît également comme la loi limite, sous certaines conditions et en régime permanent, de la probabilité pour qu'il n'y ait aucune avarie dans un intervalle (o, t), pour un processus complexe de renouvellement engendré par les avaries successives et indépendantes de n composantes d'un équipement lorsque n tend vers l'infini. De même l'intervalle entre deux avaries successives et, sous des conditions plus restrictives, l'âge de la première avarie suivent des lois exponentielles (cf. Drenick (IV)).

Mais il serait dangereux d'appliquer les tests de durée de vie en vue de déterminer les caractéristiques d'un matériel, ou d'un lot de matériels, en partant de l'hypothèse d'une loi de distribution des durées de vie exponentielle, sans s'être assuré de la validité de celle-ci (ces notions seront développées dans un prochain article).

C'est pourquoi nous signalons ci-dessous quelques tests plus ou moins connus de la forme exponentielle d'une distribution.

Ces tests imposent la mise en observation de n matériels pendant un temps donné ou variable pour un processus de renouvellement déterminé (remplacement ou non des matériels avariés pendant le temps d'observation), suivant le test considéré.

A - REALISATION PRATIQUE DES TESTS -

1/ Procédé graphique.

La fonction de répartition de (1) est de la forme :

$$F(t) = 0 \qquad \text{pour} \quad t < 0$$

$$F(t) = 1 - e^{-\frac{t}{\theta}} \qquad \text{pour} \quad t \geqslant 0 \qquad \vartheta > 0 \quad , \quad \vartheta = \frac{1}{\lambda}$$

(c'est-à-dire que ϑ désigne la durée de vie moyenne).

Donc, si l'on pose $y = \log \left(\frac{1}{(1 - F(t))}\right) = \frac{t}{\vartheta}$ et que l'on porte pour chaque valeur de t les valeurs correspondantes de y en ordonnée, on doit obtenir une droite de pente $\frac{1}{\vartheta}$.

D'où le test graphique suivant :

On place n matériels sous contrôle et soient $t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n$ les temps auxquels se produisent les avaries.

t_i est une variable aléatoire ; il en est de même de la variable $X_i = F(t_i)$ qui, pour un échantillon de n peut prendre les valeurs $F(t_1)$, $F(t_2), \ldots F(t_n)$, tirées dans une distribution uniforme sur l'intervalle (0 - 1).

Pour un matériel particulier, on a $f(t) = \frac{1}{\vartheta} e^{-\frac{t}{\theta}}$, $F(t) = 1 - e^{-\frac{t}{\theta}}$.

La probabilité que la 1ère avarie se produise dans l'intervalle $(t_i, t_i + dt_i)$ est donc $f(t_i)\,dt_i$ = Probabilité d'avoir : [i - 1 matériels avariés dans (o, t_i), un matériel avarié dans $(t_i,\ t_i + dt_i)$, n - 1 matériels non avariés].

Soit :

$$f(t_i) = \frac{n!}{(i-1)!\ 1!\ (n-i)!} \left(1 - e^{-\frac{t_i}{\theta}}\right)^{i-1} \frac{1}{\vartheta} e^{-\frac{t_i}{\theta}} \left(e^{-\frac{t_i}{\theta}}\right)^{n-i}$$

La loi de probabilité élémentaire de la variable aléatoire $u = X_i$ est donc de la forme :

$$g(u)\,du = \int_o^1 \frac{n!}{(i-1)!\ 1!\ (n-i)!} u^{i-1}(i - u)^{n-i}\,du$$

On a alors :

$$E(X_i) = \int_o^1 \frac{n!}{(i-1)!\ 1!\ (n-i)!} u^{i}(1 - u)^{n-i}\,du$$

Compte tenu de :

$$\int_o^1 u^i(1-u)^{n-i}\,du = B(i+1,\ n+1-i) = \frac{i!\,(n-i)!}{(n+1)!}$$

On a finalement :

$$E[F(t_i)] = \frac{i}{n+1}$$

On représente alors y_i en fonction de t_i. Si la loi des durées de vie est bien exponentielle, les points doivent être alignés suivant une droite passant par l'origine.

Remarques :

1/ Si la distribution des durées de vie est une loi exponentielle à deux paramètres :

$$\left[f(t\ ;\ \vartheta, A) = \frac{1}{\vartheta} e^{-\frac{(t-A)}{\theta}} \qquad 0 \leqslant A \leqslant t\right]$$

la droite d'ordonnée Log $\frac{1}{1 - F(t_i)}$ devra passer par le point $(t = A, y = 0)$.

2/ Soient N_o matériels au temps $t = 0$. Soit $N(t)$ le nombre de matériels survivants au bout d'un temps t.

Toujours dans l'hypothèse de la loi exponentielle, $N(t) = N_o\, e^{-\frac{t}{\theta}}$ et Log $N(t)$ dépend linéairement de t.

C'est une autre forme du test précédent.

2/ <u>Test du χ^2.</u>

Si on possède un nombre assez important de résultats on peut estimer ϑ $(\vartheta = \hat{\vartheta})$ puis calculer le nombre correspondant d'avaries par intervalle de temps, e_i. $e_i = n\, p_i$ avec :

$$p_1 = \int_o^{t_1} \frac{1}{\hat{\vartheta}} e^{-\frac{x}{\theta}} dx$$

$$p_i = \int_{t_{i-1}}^{t_i} \frac{1}{\hat{\vartheta}} e^{-\frac{x}{\theta}} dx \qquad i = 2 \ldots . k - 1$$

$$p_k = \int_{t_{k-1}}^{\infty} \frac{1}{\hat{\vartheta}} e^{-\frac{x}{\theta}} dx$$

Soit O_i le nombre observé d'avaries dans l'intervalle de temps i.

On forme $\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$ variable tabulée dont la loi dépend d'un paramètre ν, nombre de degrés de liberté, égal ici à $k - 1$.

3/ <u>Test faisant appel à la distribution conditionnelle de la somme des durées de vie.</u>

On sait qu'à une distribution des durées de vie de forme exponentielle correspond une loi d'apparition des avaries poissonienne :

$$\Pr\{k = r \text{ dans l'intervalle } (o,\, t)\} = e^{-\lambda t} \frac{(\lambda t)^r}{r!}$$

où λ est le taux moyen d'avaries.

<u>Propriétés découlant de l'utilisation de la loi de POISSON.</u>

Soient r évènements, à arrivée poissonienne, se produisant aux instants $\tau_1, \tau_2, \ldots, \tau_r$.

Soit (O, T) l'intervalle préassigné pendant lequel on observe ces r évènements. $(T > \tau_r)$.

On démontre que ces temps τ_i peuvent être considérés, si l'on ne tient pas compte de la relation d'ordre qui les lie, comme r observations indépendantes d'une variable aléatoire uniformément répartie sur l'intervalle (o, T).

En effet, soit t_i le temps au bout duquel se produit le $i^{\text{ème}}$ évènement.

On a alors :

$$\Pr\{\tau_1 < t_1 < \tau_1 + \Delta\tau_1,\quad \tau_2 < t_2 < \tau_2 + \Delta\tau_2,\ \ldots .\ \tau_r < t_r < \tau_r + \Delta\tau_r,$$

sachant que r évènements se produisent dans (o, T)

$$= \frac{e^{-\lambda t_1} \lambda \Delta \tau_1 e^{-\lambda (t_2 - t_1)} \lambda \Delta \tau_2 \dots e^{-\lambda (T - \tau_r)} \lambda \Delta \tau_r}{e^{-\lambda T} \frac{(\lambda T)^r}{r!}}$$

$$= \frac{r!}{T^r} \prod_{i=1}^{r} \Delta \tau_i$$

ce qui montre que les τ_i sont uniformément répartis sur (o, T) ou que les $\frac{\tau_i}{T}$ sont uniformément répartis sur (0, 1).

Pour r même assez faible, $\tau = \sum_{i=1}^{r} \tau_i$ suit approximativement une loi normale de moyenne :

$$E\left[\sum_{i=1}^{r} \tau_i\right] = \sum_{i=1}^{r} E[\tau_i] = r \frac{T}{2}$$

et de variance :

$$V\left[\sum_{i=1}^{r} \tau_i\right] = \sum_{i=1}^{r} V(\tau_i) = r \frac{T^2}{12}$$

La variable $\dfrac{\tau - r \frac{T}{2}}{T \sqrt{\frac{r}{12}}}$ suit donc une loi normale centrée réduite ce qui permet de tester la validité de l'hypothèse faite (loi de distribution poissonienne des avaries).

Supposons maintenant que l'on arrête l'observation lorsque le $r^{ième}$ évènement se produit, c'est-à-dire au temps τ_r.

On démontre, de façon analogue à ce qui précède, que les (r - 1) quantités $\tau_1 \dots \tau_{r-1}$ peuvent être considérées, si l'on ne tient pas compte de leur relation d'ordre, comme (r - 1) observations indépendantes d'une variable aléatoire uniformément répartie sur l'intervalle (o, τ_r).

Pour r pas trop faible, $\tau = \sum_{i=1}^{r=1} \tau_i$ suit approximativement une loi normale de moyenne $\frac{(r-1)\tau_r}{2}$ et de variance $\frac{(r-1)\tau_r^2}{12}$.

On forme $\dfrac{\tau - \frac{(r-1)\tau_r}{2}}{\tau_r \sqrt{\frac{r-1}{12}}}$ variable normale centrée réduite. Comme précédemment ceci permet de tester si la distribution des avaries est de POISSON.

Tout ce qui précède s'applique au cas où les matériels avariés sont remplacés immédiatement par des nouveaux en état de fonctionner.

Si l'on est dans le cas où l'on ne remplace pas les matériels avariés, on considère, au lieu des t_i, les sommes des durées de vie observées avant l'apparition de la $i^{ème}$ avarie : $T(\tau_i)$ $i = 1 \dots r$.

$$T(\tau_1) \leqslant T(\tau_2) \leqslant T(\tau_3) \leqslant \dots \leqslant T(\tau_r)$$

Dans le cas où le test se termine lorsque la somme des durées de vie égale T^*, valeur préassignée, le nombre d'avaries observées r est une variable aléatoire avec :

$$
\begin{aligned}
T(\tau_1) &= n\,\tau_1 \\
T(\tau_2) &= \tau_1 + (n-1)\tau_2 \\
&\vdots \\
T(\tau_r) &= \tau_1 + \tau_2 + \ldots + \tau_{r-1} + (n-r+1)\tau_r
\end{aligned}
$$

On démontre, comme précédemment, que les sommes des durées de vie $T(\tau_1) \ldots T(\tau_r)$ peuvent être considérées comme r observations indépendantes d'une variable aléatoire uniformément répartie sur $[0,\ T^*]$.

Dans le cas où le test se termine dès l'apparition de la $r^{\text{ième}}$ avarie, on démontre que les $T(\tau_i)$ $(i = 1 \ldots r-1)$ peuvent être considérées comme $r-1$ observations indépendantes d'une variable aléatoire uniformément répartie sur $[0,\ T(\tau_r)]$.

Dans les deux cas on testera par un χ^2 l'ajustement des distributions observée et théorique des $T(\tau_i)$.

D'autre part, cette propriété de la distribution conditionnelle des $T(\tau_i)$ fournit un moyen pratique de se rendre compte si le taux d'avarie est vraiment constant (comme il devrait l'être dans le cas d'une loi exponentielle). Les tests 4, 5, 6, 7, 8 en sont des illustrations.

4/ Test des avaries se produisant anormalement tôt.

Le caractère uniforme de la répartition des $T(\tau_i)$ exclut la possibilité d'un nombre important d'avaries se produisant très tôt.

Or, il se peut que même si la distribution de base est exponentielle, il se produise de ces avaries anormalement précoces. Comment tester cette sorte de déviation de la loi exponentielle ?

Soient $\tau_1 \leqslant \tau_2 \leqslant \ldots \leqslant \tau_r$ les temps auxquels se sont produites les r premières avaries. On veut tester si τ_1 est anormalement petit.

D'après ce qui précède, $T(\tau_1)$ somme des durées de vie sur l'intervalle $(o,\ \tau_1)$ et $T(\tau_r - \tau_1)$ somme des durées de vie sur l'intervalle $(\tau_1,\ \tau_r)$ sont distribuées indépendamment l'une de l'autre.

On démontre également que :

$\dfrac{2\,T(\tau_1)}{\vartheta}$ suit une loi du χ^2 à 2 degrés de liberté (on fait $r = 1$)

$\dfrac{2\,T(\tau_r - \tau_1)}{\vartheta}$ suit une loi du χ^2 à $2(r-1)$ degrés de liberté (on fait $r = r - 1$)

D'où $\boxed{\dfrac{(r-1)\,T(\tau_1)}{T(\tau_r - \tau_1)} = F}$ suit une loi de FISHER-SNEDECOR à 2 et $2r - 2$ degrés de liberté.

τ_1 sera considéré comme anormalement petit si le F est trop faible. Plus précisément, si α désigne le seuil de confiance du test, on dira que τ_1 est anormalement petit si :

$$T(\tau_1) < F_\alpha^* \frac{T(\tau_r - \tau_1)}{r - 1} \quad \text{où } F_\alpha^* \text{ est défini par :}$$

$$\Pr\{F(2,\ 2r - 2) \leqslant F_\alpha^*\} = \alpha$$

De façon analogue on obtient un critère permettant de tester si à la fois τ_1 et τ_2 sont anormalement petits. Pour cela on montre que :

$$\boxed{\frac{(r - 2)\ T(\tau_2)}{2\ T(\tau_r - \tau_2)}}$$ est distribué suivant une loi de FISHER-SNEDECOR

à 4 et 2r - 4 degrés de liberté et la condition, telle que si elle est réalisée τ_1 et τ_2 seront considérés comme anormalement petits, s'écrit :

$$T(\tau_2) < F_\alpha^{**} \frac{2\ T(\tau_r - \tau_2)}{r - 2}$$

avec :

$$\Pr\{F(4,\ 2r - 4) \leqslant F_\alpha^{**}\} = \alpha$$

5/ <u>Test d'un temps d'apparition de la première avarie anormalement long</u>.

Si la distribution de base est bien exponentielle mais à deux paramètres :

$$\left[f(t\ ;\ \vartheta) = \frac{1}{\vartheta}\ e^{-\frac{(t - A)}{\theta}} \quad A > 0 \right]$$ (A désignant la durée de vie minimum)

le temps d'apparition de la première avarie sera anormalement long (si l'on a fait l'hypothèse d'une distribution de base exponentielle à un seul paramètre).

Le test utilisé est basé sur la distribution du rapport :

$$\boxed{\frac{(r - 1)\ T(\tau_1)}{T(\tau_r - \tau_1)} = F}$$

On a vu que pour A = 0, F suivait une loi de FISHER-SNEDECOR à 2 et 2r - 2 degrés de liberté.

On rejettera donc l'hypothèse A = 0 (au profit de A > 0) si :

$$T(\tau_1) > G_\alpha^* \frac{T(\tau_r - \tau_1)}{r - 1}$$

où G_α^* est défini par : $\Pr\{F(2,\ 2r - 2) \geqslant G_\alpha^*\} = \alpha$

6/ <u>Test du caractère significatif de la différence entre la durée de vie moyenne</u> (ou taux d'avarie moyen) <u>pendant la première moitié du test et la durée de vie moyenne</u> (ou taux d'avarie moyen) <u>pendant la seconde moitié</u>.

Ceci est un procédé grossier mais rapide de détecter les déviations importantes pouvant se produire avec le temps. Le choix des deux périodes de comparaison dépend de considérations à priori.

Soient $\tau_1\, \tau_2 \ldots \tau_{2r}$ les temps d'apparition des 2r premières avaries. Comparons ce qui se passe dans les intervalles (o, τ_r) et (τ_r, τ_{2r}).

Si la distribution reste exponentielle avec une durée de vie moyenne ϑ fixe, $T(\tau_r)$ et $T[\tau_{2r} - \tau_r]$ sont indépendants l'un de l'autre.

$$\frac{2\,T(\tau_r)}{\vartheta} \text{ suit une loi du } \chi^2 \text{ à 2r degrés de liberté,}$$

$$\frac{2\,T(\tau_{2r} - \tau_r)}{\vartheta} \text{ également.}$$

D'où :

$$\frac{T(\tau_r)}{T(\tau_{2r} - \tau_r)} = F(2r,\, 2r)$$

ce qui permet de tester l'hypothèse.

7/ Test de la constance de la durée de vie moyenne.

Ce test est une généralisation du précédent. Soient kr temps d'apparition des avaries :

$$\tau_{11} \leqslant \tau_{12} \leqslant \ldots \leqslant \tau_{1r} \leqslant \tau_{21} \leqslant \ldots \leqslant \tau_{2r} \leqslant \ldots \leqslant \tau_{k1} \leqslant \ldots \leqslant \tau_{kr}$$

rangés en k groupes de dimensions r.

On suppose que ϑ est constant à l'intérieur de chaque groupe et on veut tester si ϑ fluctue d'un groupe à l'autre.

On sait que :

$$\frac{2\,T[\tau_{jr} - \tau_{(j-1)r}]}{\vartheta} = \chi^2(2r) \qquad \text{pour} \quad j = 1 \ldots k$$

d'où :

$$T(\tau_{kr}) = T[\tau_{1r}] + T[\tau_{2r} - \tau_{1r}] + \ldots + T[\tau_{kr} - \tau_{(k-1)r}]$$

est tel que :

$$\frac{2\,T[\tau_{kr}]}{\vartheta} = \chi^2(2\,kr)$$

ou encore, d'après le test d'homogénéité des variances de Bartlett :

$$\frac{2rk\left\{\mathrm{Log}\,\dfrac{T(\tau_{kr})}{k} - \dfrac{1}{k}\left[\mathrm{Log}\,T(\tau_{1r}) + \mathrm{Log}\,T(\tau_{2r} - \tau_{1r}) + \ldots + \mathrm{Log}\,T[\tau_{kr} - \tau_{(k-1)r}]\right]\right\}}{1 + \dfrac{k+1}{6\,rk}} \sim \chi^2(k-1)$$

L'hypothèse que ϑ est constant d'un groupe à l'autre sera rejetée si le terme de gauche de l'équation est supérieur à $\chi^2_\alpha(k-1)$ défini par :

$$\Pr\{\chi^2(k-1) > \chi^2_\alpha(k-1)\} = \alpha = \text{seuil de confiance du test.}$$

8/ Cas particulier : r = 1.

Supposons que dans le test précédent nous fassions r = 1 :

$$\tau_{11} = \tau_1, \ldots, \tau_{k1} = \tau_k$$

La relation précédente s'écrit alors :

$$\frac{2k\left\{\mathrm{Log}\,\frac{T(\tau_k)}{k} - \frac{1}{k}\left[\mathrm{Log}\,T(\tau_1) + \mathrm{Log}\,T(\tau_2 - \tau_1) + \ldots + \mathrm{Log}(\tau_k - \tau_{k-1})\right]\right\}}{1 + \frac{k+1}{6k}} \sim \chi^2(k-1)$$

Remarques :

a) Le numérateur du membre de gauche se réduit à :

$$-2\sum_{j=1}^{k} \mathrm{Log}\,\frac{T(\tau_j - \tau_{j-1})}{\overline{T(\tau_j - \tau_{j-1})}}$$

où :

$$\overline{T(\tau_j - \tau_{j-1})} = \sum_{j=1}^{k} \frac{T(\tau_j - \tau_{j-1})}{k} = \frac{T(\tau_k)}{k}$$

b) Il est intéressant de noter que le test du rapport de vraisemblance consistant à tester l'hypothèse H_0 (distribution des sommes des durées de vie entre avaries successives de la forme : $f(x\,;\vartheta) = \frac{1}{\vartheta}\,e^{-x/\theta}$) contre l'hypothèse H_1 (densité de probabilité de la forme :

$$f(x\,;\vartheta,\beta) = \frac{1}{\Gamma(\beta)\,\vartheta^{\beta}}\,x^{\beta-1}\,e^{-x/\theta} \qquad \beta > 0)$$

conduit également à la relation précédente.

c) Ce test est une autre formulation du test 2, basé sur l'emploi du χ^2 ainsi que du test 3, basé sur l'uniformité de la distribution conditionnelle des sommes de durées de vie.

A ce propos revenons un moment au test 3 pour le mettre en parallèle avec les deux tests qui précèdent.

Il est clair que si u est une variable aléatoire uniformément répartie sur (0, 1), - 2 Log u suit une loi du χ^2 à 2 degrés de liberté.

Donc $-2\sum_{i=1}^{r} \mathrm{Log}\,\frac{T(\tau_i)}{T^*}$ est distribué suivant une loi du χ^2 à 2r degrés de liberté, où T^* est une somme de durée de vie préassignée et le nombre d'avaries, r, une variable aléatoire.

De la même manière, dans le cas où le test est arrêté après r avaries, $-2\sum_{i=1}^{r-1} \mathrm{Log}\,\frac{T(\tau_i)}{T(\tau_r)}$ est distribué comme un χ^2 à 2r - 2 degrés de liberté.

d) Relation entre les tests 7 et 8 et le test L_1 de Neyman et Pearson.

d_1) Test L_1 de Neyman et Pearson.

On tire respectivement de k populations normales, un échantillon de taille n_i (i = 1 .... k).

On désire tester l'hypothèse que les k populations ont même variance les moyennes μ_i étant quelconques.

Soit : $\bar{x}_i = \sum_{j=1}^{n_i} \frac{x_{ij}}{n_i}$ la moyenne empirique du ième échantillon.

et : $s_i^2 = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2/n_i$ sa variance empirique.

On définit :

$$s_a^2 = \sum_{i=1}^{k} \frac{n_i\, s_i^2}{N} \quad \text{où} \quad N = \sum_{i=1}^{k} n_i$$

La fonction de vraisemblance peut s'écrire :

$$p = \frac{1}{(2\pi)^{\frac{N}{2}} \prod_{i=1}^{k} \sigma_i^{n_i}} \exp\left\{- \sum_{i=1}^{k} n_i \frac{(\bar{x}_i - \mu_i)^2 + s_i^2}{2\sigma_i^2}\right\}$$

d'où l'on tire les estimations suivantes :

$$\mu_i = \bar{x}_i$$

$$\sigma_i^2 = \begin{cases} s_a^2 & \text{pour l'hypothèse nulle} \\ s_i^2 & \text{pour l'hypothèse alternative} \end{cases}$$

et :

$$\lambda_{H_0} = \frac{\text{fonction de vraisemblance dans l'hypothèse alternative}}{\text{fonction de vraisemblance dans l'hypothèse nulle}}$$

$$\lambda_{H_0} = \prod_{i=1}^{k} \left(\frac{s_i^2}{s_a^2}\right)^{\frac{n_i}{2}}$$

Le rejet de l'hypothèse H_0 a lieu pour les petites valeurs de λ_{H_0}.

On démontre que λ_{H_0} suit approximativement, pour n grand, une loi du χ^2 à (k - 1) degrés de liberté [on part de la loi connue de l'ensemble $(s_1^2 \ldots s_k^2)$].

On considère la quantité $(\lambda_{H_0})^{2/N}$ - Dans le cas particulier où tous les n_i sont égaux à n, $(\lambda_{H_0})^{2/N}$ se réduit à la fonction L_1 de Neyman et Pearson.

$$\boxed{L_1 = \prod_{i=1}^{k} \left(\frac{s_i^2}{s_a^2}\right)^{1/k}} \qquad (N = kn)$$

Nayer a calculé les valeurs de L_1 pour des seuils de confiance de 1 % et 5 %, des n égaux à 2, 5, 20, 50 et des k égaux à 2, 3, 4, 5, 10, 20, 25, 50.

d_2) Relation avec les tests 7 et 8.

On a vu que $\frac{2\,T(\tau_{jr} - \tau_{(j-1)r})}{\vartheta}$ suivait une loi du χ^2 à 2r degrés de liberté. On peut donc énoncer le test 7 sous la forme suivante :

On extrait des échantillons de taille $n = 2r + 1$ de k populations normales et on désire tester l'égalité des variances de ces populations (les moyennes étant quelconques). Ici s_i^2 est remplacé par $T(\tau_{ir} - \tau_{(i-1)r})$ et s_a^2 par $\sum_{i=1}^{k} \frac{T(\tau_{ir} - \tau_{(i-1)r})}{k}$. L_1 est alors égal à :

$$\boxed{L_1 = \prod_{i=1}^{k} \left[\frac{T(\tau_{ir} - \tau_{(i-1)r})}{\frac{T(\tau_{kr})}{k}} \right]^{1/k}}$$

Bartlett a montré qu'un test meilleur, car sans distorsion, consistait à définir $\mu^{2/\nu}$ tel que :

$$\mu^{2/\nu} = \prod_{i=1}^{k} \left[\frac{\nu s_i^2}{\sum_{i=1}^{k} \nu_i s_i^2} \right]^{\nu_i/\nu} \quad \text{où} \quad \nu_i = n_i - 1 \quad \text{et} \quad \nu = \sum_{i=1}^{k} \nu_i$$

puis :

$$\frac{-2 \operatorname{Log} \mu}{c} \quad \text{où} \quad c = 1 + \frac{1}{3(k-1)} \left\{ \sum_{i=1}^{k} \frac{1}{\nu_i} - \frac{1}{\nu} \right\}$$

quantité suivant approximativement une loi du χ^2 à $k - 1$ degrés de liberté.

Dans le test de durée de vie envisagé,

$$\begin{cases} n_i = n = 2r + 1 \\ \nu_i = 2r \end{cases} \quad \text{pour tout } i$$

et :

$$\nu = \sum_{i=1}^{k} \nu_i = 2\,rk$$

d'où :

$$c = 1 + \frac{1}{3(k-1)} \left\{ \frac{k}{2r} - \frac{1}{2rk} \right\} = 1 + \frac{k+1}{6rk}$$

de plus :

$$\mu^{2/\nu} = \mu^{1/rk} = L_1$$

d'où :

$$\frac{-2 \operatorname{Log} \mu}{c} = \frac{-2\,rk \operatorname{Log} L_1}{c} \sim \chi^2(k-1)$$

on retrouve bien la formule du test 7.

Dans le cas où $r = 1$ (test 8) :

$$\frac{-2k \operatorname{Log} L_1}{1 + \frac{k+1}{6k}} \sim \chi^2(k-1)$$

avec : $$\boxed{L_1 = \prod_{i=1}^{k} \left[\frac{T(\tau_i - \tau_{i-1})}{\frac{T(\tau_k)}{k}} \right]^{1/k}}$$

on retrouve bien la formule du test 8.

9/ Test basé sur la distribution du F maximum.

Ce test est dû à Hartley.

Les hypothèses sont les mêmes que celles du test L_1. En particulier on désire toujours tester l'égalité des variances des k populations.

s_i^2 est pris comme estimation de σ_i^2 avec ν degrés de liberté.

On forme $F_{max} = \frac{\max_i s_i^2}{\min_i s_i^2}$ et on rejette l'hypothèse d'égalité des variances si le F_{max} est trop grand.

En termes de test de durée de vie, cela peut se mettre sous la forme suivante :

Supposons que l'on ait k sommes de durées de vie $\{T_i\}$, $i = 1 \ldots k$, chacune d'entre elles étant formée après les r premières avaries, d'un échantillon de taille n et supposons que l'on désire tester la constance de ϑ entre échantillons.

Le F_{max} devient :

$$\boxed{F_{max} = \frac{\max T_i}{\min T_i}}$$

et on entre dans les tables avec un nombre d'échantillons égal à k et un nombre de degrés de liberté pour chaque échantillon égal à $\nu = 2r$.

10/ Test sur les périodes anormalement longues sans avaries.

Le test 5 traitait du cas où la première avarie se produisait après un temps anormalement long. Celui-ci permet de tester n'importe quel intervalle.

soit : $$\boxed{Z = \frac{\max\{T(\tau_1),\ T(\tau_2 - \tau_1) \ldots\ T(\tau_n - \tau_{n-1})\}}{T(\tau_n)}}$$

où $\tau_1 \leqslant \tau_2 \leqslant \ldots \leqslant \tau_n$ sont les n premiers temps ordonnés d'apparition des avaries.

Fisher a montré que :

$$\Pr\{Z > Z_0\} = \sum_{k=1}^{r} C_n^k (-1)^{k-1} (1 - k Z_0)^{n-1} \tag{1}$$

où r est le plus grand nombre entier inférieur à $\frac{1}{Z_0}$.

(1) est la distribution de $\max_{1 \leqslant i \leqslant n} \{x_i\} / \sum_{i=1}^{n} x_i$ où $x_1, x_2 \ldots x_n$ sont n observations indépendantes d'une variable aléatoire X de densité de probabilité $\frac{1}{\vartheta} e^{-x/\theta}$ $(x > 0,\ \vartheta > 0)$.

Dans le test : $x_1 = \frac{2\,T(\tau_1)}{\vartheta} \ldots\ldots x_n = \frac{2\,T(\tau_n - \tau_{n-1})}{\vartheta}$. Fisher donne les valeurs de Z_0 telles que $\Pr\{Z > Z_0\} \leqslant \alpha$.

pour : $\alpha = 0,05 - 0,01$ et $n \leqslant 50$.

Généralisation due à Cochran.

Supposons qu'on tire des échantillons d'une population normale de variance σ^2. On obtient k estimations indépendantes de σ^2, $s_1^2 \ldots\ldots s_k^2$, chacune avec ν degrés de liberté.

Cochran définit la statistique g_i par :

$$g_i = \frac{s_i^2}{\sum_{i=1}^{k} s_i^2} \qquad i = 1, 2, \ldots\ldots k$$

et donne la distribution de $\max_i \{g_i\}$.

Puisque $\frac{\nu\, s_i^2}{\sigma^2}$ est distribué comme un χ^2 à ν degrés de liberté, les résultats de Cochran s'appliquent au cas particulier où l'on étudie la distribution de $\max_{1 \leqslant i \leqslant k} \{x_i\} / \sum_{i=1}^{k} x_i$ où $x_i \ldots\ldots x_k$ sont k observations indépendantes d'une variable aléatoire suivant une loi du χ^2 à 2r degrés de liberté. Dans le test on peut identifier les x_i aux quantités :

$$x_1 = \frac{2\,T(\tau_{1r})}{\vartheta}, \quad x_2 = \frac{2\,T(\tau_{2r} - \tau_{1r})}{\vartheta}, \ldots\ldots x_k = \frac{2\,T(\tau_{kr} - \tau_{(k-1)r})}{\vartheta}$$

et :

$$\boxed{g_i = \frac{T(\tau_{ir} - \tau_{(i-1)r})}{T(\tau_{kr})}} \qquad i = 1, 2, \ldots\ldots k$$

On retrouve le test de Fisher précédemment défini en faisant $k = n$ et $\nu = 2$.

11/ Méthode graphique basée sur le test de Kolmogorov-Smirnov.

a) Considérons un test arrêté quand la somme des durées de vie a atteint une certaine valeur T^* préassignée.

Le nombre d'avaries observées, r, est une variable aléatoire et soient $\tau_1 \leqslant \tau_2 \leqslant \ldots\ldots \leqslant \tau_r$ les temps d'apparition des avaries et $T(\tau_1) \leqslant T(\tau_2) \leqslant \ldots\ldots \leqslant T(\tau_r)$ les sommes de durées de vie correspondantes.

On a vu précédemment que les $T(\tau_i)$ étaient réparties uniformément sur $[0, T^*]$ si la distribution de base est exponentielle, de paramètre constant.

D'après Barnard et Cox, une méthode graphique de tester cette hypothèse est de représenter la proportion d'avaries (relative à r) qui se produisent en un temps $T(\tau)$, en fonction de $T(\tau)$. Si l'hypothèse testée est vraie le nuage de points doit être sensiblement aligné le long de la droite joignant l'origine au point $(T^*, 1)$.

En se servant des résultats de Kolmogorov et de Smirnov on peut établir un test de signification sur la dimension de $r^{1/2} D_r$, où D_r est la différence maximum (mesurée verticalement) entre le point observé et le point de la ligne de tendance.

Expression analytique.

On définit la quantité aléatoire $F_r[T(\tau)]$ par : $r\,F_r[T(\tau)]$ = nombre d'avaries se produisant pendant un temps $T(\tau)$.

La fonction de répartition théorique de $T(\tau)$ est :

$$F[T(\tau)] = \frac{T(\tau)}{T^*} \qquad \text{pour} \qquad 0 \leqslant T(\tau) \leqslant T^*$$

Soit, si D_r est la statistique de Kolmogorov-Smirnov :

$$r\,D_r = r \max_{0<T(\tau)<T^*} \left| F_r[T(\tau)] - F[T(\tau)] \right|$$

$$r\,D_r = \max_{0<T(\tau)<T^*} \left| \begin{matrix}\text{nombre d'avaries}\\ \text{se produisant en}\\ \text{un temps } T(\tau)\end{matrix} - \frac{r\,T(\tau)}{T^*} \right|$$

Si r est grand ($\geqslant 80$), l'hypothèse que la distribution est exponentielle de paramètre constant sera rejetée avec un seuil de confiance de 10 % si la différence maximum mesurée par $r^{1/2} D_r$ dépasse 1,22. Les valeurs de $r^{1/2} D_r$ à 5 % et 1 % sont respectivement 1,36 et 1,63. Pour de petites valeurs de r on utilise les tables de Z.W. Birnbaum et Massey.

b) Supposons maintenant que l'on arrête le test après l'apparition de la $(n+1)^{\text{ième}}$ avarie.

On a vu que, toujours sous la même hypothèse, $T(\tau_1) \ldots . T(\tau_n)$ étaient distribués uniformément sur $[0, T(\tau_{n+1})]$.

Un graphique analogue au précédent peut être tracé en remplaçant r par n et T^* par $T(\tau_{n+1})$.

Le test de signification sera basé sur $n^{1/2} D_n$ où D_n mesure la différence verticale maximum entre le point observé et le point de la droite joignant l'origine au point $(T(\tau_{n+1}), 1)$.

Expression analytique.

On définit la quantité aléatoire $F_n[T(\tau)]$ par : $n\,F_n[T(\tau)]$ = nombre d'avaries se produisant pendant un temps $T(\tau)$.

La fonction de répartition théorique de $T(\tau)$ est :

$$F[T(\tau)] = \frac{T(\tau)}{T(\tau_{n+1})} \qquad \text{pour} \qquad 0 \leqslant T(\tau) \leqslant T(\tau_{n+1})$$

Soit, si D_n est la statistique de Kolmogorov-Smirnov :

$$n\,D_n = n \max_{0<T(\tau)<T(\tau_{n+1})} \left| F_n[T(\tau) - F\ T(\tau)] \right|$$

$$= \max_{0<T(\tau)<T(\tau_{n+1})} \left| \begin{matrix}\text{nombre d'avaries}\\ \text{se produisant en}\\ \text{un temps } T(\tau)\end{matrix} - \frac{n\,T(\tau)}{T(\tau_{n+1})} \right|$$

12/ Test basé sur le taux d'avarie conditionnel.

Une propriété caractéristique de la distribution exponentielle est que la probabilité conditionnelle d'avarie d'un matériel dans l'intervalle $(t, t + \Delta t)$ sachant qu'il est encore en vie à l'instant t, est indépendante de t.

Conséquence pratique.

Soit N un grand nombre de matériels à tester. On divise l'axe des temps en intervalles égaux à T : (0, T), (T, 2 T)....

Soient n_1, n_2, n_3,... les nombres de matériels avariés dans ces intervalles de temps.

Si l'hypothèse de base est réalisée, $\frac{n_1}{N}, \frac{n_2}{N - n_1}, \ldots\ldots, \frac{n_k}{N - n_1 \ldots n_{k-1}}$ doivent être sensiblement constants et égaux au taux d'avarie (relatif à un intervalle T).

Si la distribution des durées de vie n'est pas exponentielle $\frac{f(t)}{1 - F(t)}$ sera quelque fonction g(t) et comme $f(t) = \frac{d\,F(t)}{dt}$.

on aura :

$$F(t) = 1 - e^{-G(t)} \qquad \text{avec} \qquad G(t) = \int_0^t g(\tau)\, d\tau$$

Cas particulier.

$$G(t) = \alpha t^k$$

La distribution de la probabilité conditionnelle d'avarie est donc théoriquement de la forme $C\, t^{k-1}$.

Les points $\left(\operatorname{Log}\frac{n_1}{N}, \operatorname{Log}\frac{T}{2}\right) \ldots \left(\operatorname{Log}\frac{n_k}{N - n_1 \ldots n_{k-1}}, \operatorname{Log}\frac{(2k-1)T}{2}\right)$ devront donc être sensiblement alignés.

13/ Test basé sur les sommes des durées de vie entre avaries successives.

Sherman et plus récemment Bartholomew ont montré que si $t_1, t_2, \ldots, t_n$ désignent les intervalles entre les arrivées poissoniennes d'évènements successifs, la statistique :

$$\varpi = \frac{\sum_{i=1}^{n} |t_i - \bar{t}|}{2 n \bar{t}} \qquad \text{où} \qquad \bar{t} = \frac{\sum_{i=1}^{n} t_i}{n}$$

peut être utilisée pour tester l'hypothèse faite sur la distribution des arrivées.

La loi de ϖ est en effet la suivante :

$$F(x) = \operatorname{Prob}\{\varpi \leqslant x\}$$

$$= 1 + \sum_{q=0}^{n-\nu-2} \sum_{p=0}^{q} C_{n-1}^{p}\, C_n^{q+1}\, C_{n+q-p-1}^{n-1} \left(\frac{n+2}{n+q-1}\right)^q \left(\frac{n-q-1}{n} - x\right)^{n-p-1}$$

où v désigne le nombre entier satisfaisant à la relation :

$$\frac{v}{n} \leqslant x \leqslant \frac{v+1}{n}$$

On démontre que pour $n \leqslant 10$, la loi de la variable $X = \frac{n\bar{\omega}}{n-1}$ peut être approximée par une loi de Pearson du type I.

Il en résulte que la variable $W = \frac{\nu_2 X}{\nu_1 (1 - X)}$ suit une loi de Fisher-Snedecor à (ν_1, ν_2) degrés de liberté. Bartholomew a tabulé les différentes valeurs de (ν_1, ν_2) en fonction de n.

On peut donc déterminer l'intervalle de confiance de ϖ par l'intermédiaire de celui de W que l'on peut plus facilement obtenir, la loi de Fisher-Snedecor étant tabulée.

Application au test de la distribution des durées de vie :

Supposons que l'observation continue jusqu'à l'obtention de n avaries. Soient $\tau_1 \leqslant \tau_2 \leqslant \ldots \leqslant \tau_n$ les instants auxquels elles se produisent et soient $T(\tau_1) \ldots . T(\tau_n - \tau_{n-1})$ les sommes des durées de vie associées.

On a alors :

$$\boxed{\bar{\omega} = \frac{\sum_{i=1}^{n} \left| T(\tau_i - \tau_{i-1}) - \frac{T(\tau_n)}{n} \right|}{2\,T(\tau_n)}}$$

Bartholomew a montré que si le test basé sur $\bar{\omega}$ n'est pas le plus efficace pour certaines alternatives, il est très bon pour un nombre important de celles-ci.

Ce test sera particulièrement utile dans le cas où l'on a aucune idée à priori sur la forme possible de l'alternative (l'hypothèse nulle étant toujours la loi exponentielle de paramètre constant).

B - COMMENTAIRES SUR L'EFFICACITE DE CES TESTS -

- Tests 1, 2, 12 : des échantillons assez substantiels ($n \geqslant 50$) sont nécessaires pour permettre de détecter les déviations importantes de la loi exponentielle.

- Tests 3, 11 : ces deux tests sont un peu plus sensibles aux changements de distribution de base. On peut encore les utiliser pour $n \geqslant 10$, cependant il est reconnu que pour de faibles tailles d'échantillons l'efficacité du test est faible.

Ils reposent tous deux sur la même propriété découlant de l'emploi d'un procédé poissonien. Toutefois il est difficile de comparer l'efficacité du test 3 à celle du 11, puisque l'efficacité dépend de l'alternative contre laquelle est testée l'hypothèse nulle (ici : fonction exponentielle de durée de vie moyenne constante).

Le test 11 a l'avantage de faire appel à une méthode graphique.

Cox a montré qu'il existe toutefois un cas où le test 3 est plus efficace que le test 11. C'est celui où on teste l'hypothèse que les apparitions des avaries successives sont poissoniennes de taux constant λ contre l'hypothèse d'un taux d'avarie dépendant du temps et de la forme $\lambda(t) = \alpha e^{\beta t}$.

- Tests 7, 8 : Ce sont des conséquences directes du test d'homogénéité des variances de Bartlett.

Moran a montré que le test 8 est asymptotiquement un des tests les plus efficaces testant H_0 (sommes des durées de vie entre deux avaries successives distribuées suivant $f(x, \vartheta) = - e^{-x/\theta}\ x > 0,\ \vartheta > 0$) contre H_1 (sommes des durées de vie entre deux avaries successives distribuées suivant une loi gamma du type III :

$$f(x, \vartheta, \beta) = \frac{1}{\Gamma(\beta)\, \vartheta^{\beta}}\, x^{\beta-1}\, e^{-x/\theta}\quad x > 0,\ \vartheta > 0).$$

C - APPLICATION NUMERIQUE -

49 matériels ont été placés sous test et les temps d'apparition des avaries τ obtenus sont les suivants : [III]

1,2	13,7	38,9	72,4	102,8	151,6	203,0
2,2	15,1	47,9	73,6	108,5	152,6	204,3
4,9	15,2	48,4	76,8	128,7	164,2	229,5
5,0	23,9	49,3	83,8	133,6	166,8	253,1
6,8	24,3	53,2	95,1	144,1	178,6	304,1
7,0	25,1	55,6	97,9	147,6	185,2	341,7
12,1	35,8	62,7	99,6	150,6	187,1	354,4

1/ Test 1.

$$y_i = \text{Log}\,\frac{n+1}{n+1-i}\qquad i = 1, \ldots\ldots, n$$

Tableau des $\dfrac{n+1}{n+1-i}$

1,020	1,190	1,428	1,785	2,380	3,571	7,142
1,041	1,219	1,470	1,851	2,500	3,846	8,333
1,063	1,250	1,515	1,923	2,631	4,166	10,000
1,086	1,282	1,562	2,000	2,777	4,545	12,500
1,111	1,315	1,612	2,083	2,941	5,000	16,667
1,136	1,351	1,666	2,173	3,125	5,555	25,000
1,162	1,388	1,724	2,272	3,333	6,250	50,000

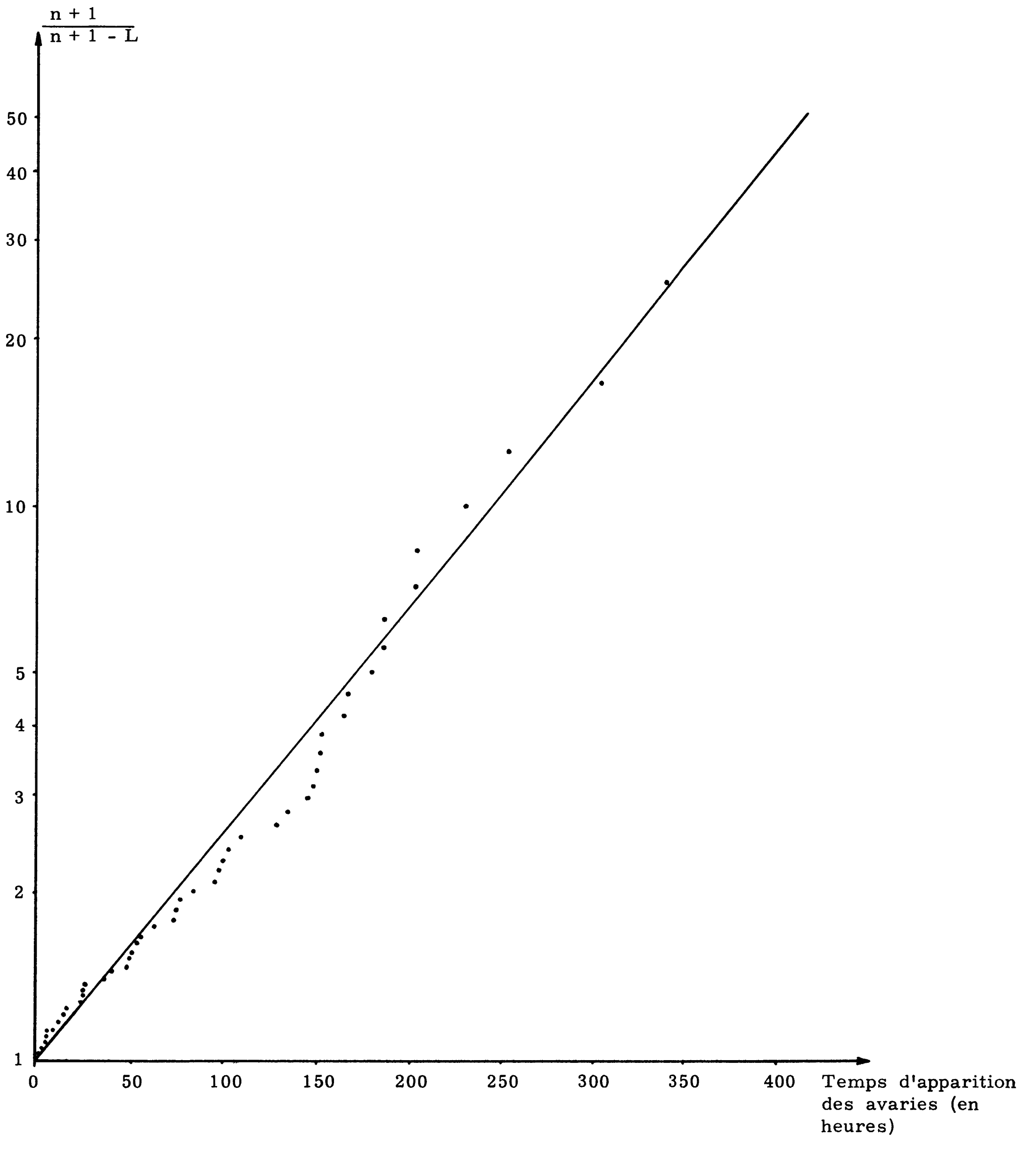
On a représenté y_i en fonction de τ_i. On constate qu'ils fluctuent autour d'une droite passant par l'origine (figure 1).

L'hypothèse semble donc réalisée.

2/ Test 2.

On estime ϑ par :

$$\hat{\vartheta} = \frac{\sum_{i=1}^{n} \tau_i}{n} = \frac{5139,6}{49} = 104,89$$



n + 1
n + 1 - L
50
40
30
20
10
5
4
3
2
1
0
50
100
150
200
250
300
350
400
Temps d'apparition des avaries (en heures)

Fig. 1

Soient les intervalles de temps : (0,20) - (20,40) -, ... , - (240,260) - > 260 .

Le nombre théorique d'avaries par intervalle de temps s'écrit : e_i = npi avec :

$$p_1 = \int_0^{t_1} \frac{1}{\hat{\vartheta}} e^{-\frac{x}{\theta}} dx = 1 - e^{-\frac{t_1}{\theta}}$$

$$p_i = \int_{t_{i-1}}^{t_i} \frac{1}{\hat{\vartheta}} e^{-\frac{x}{\theta}} dx = e^{-\frac{t_{i-1}}{\theta}} - e^{-\frac{t_i}{\theta}} \qquad i = 2, \ldots , (k-1)$$

$$p = \int_{t_k}^{\infty} \frac{1}{\hat{\vartheta}} e^{-\frac{x}{\theta}} dx = e^{-\frac{t_k}{\theta}}$$

si t_i désigne la limite supérieure du ième intervalle.

On trouve :

$t_{i-1} - t_i$	e_i	o_i
0 - 20	8,52	10
20 - 40	7,01	5
40 - 60	5,82	5
60 - 80	4,78	4
80 - 100	3,98	4
100 - 120	3,28	2
120 - 140	2,71	2
160 - 180	1,86	3
180 - 200	1,52	2
200 - 220	1,27	2
2 0 - 240	1,05	1
240 - 260	0,86	1
260	4,11	3
	49	49

les o_i sont les nombres observés d'avaries par intervalle de temps.

le χ^2 correspondant à 13 degrés de liberté vaut : 7,042.

or :

$$\Pr\{\chi^2 > 7,042\} \# 0,90$$

donc la loi peut être considérée comme exponentielle.

3/ Test 3.

On peut calculer les $T(\tau_i) = \tau_1 + \tau_2 + , \ldots , + \tau_{i-1} + (n - i + 1)\tau_i$. On trouve :

58,8	614,6	1553,8	2575,5	3306,3	4185,8	4670,5
106,8	672,0	1859,8	2607,9	3420,3	4198,8	4678,3
233,7	676,0	1876,3	2691,1	3804,1	4338,0	4804,3
238,3	1015,3	1905,1	2866,1	3892,3	4366,6	4898,7
319,3	1030,5	2026,0	3137,3	4070,8	4484,6	5051,7
328,1	1060,1	2098,0	3201,7	4126,8	4544,0	5126,9
547,4	1445,3	2303,9	3239,1	4171,8	4559,2	5139,6

On sait que les 48 premières sommes de durées de vie sont réparties uniformément sur [0,5139,6]. On va tester par un χ^2 si cette hypothèse est réalisée.

L'intervalle [0,5139,6] a été divisé en 10 parties égales et les nombres observés et théoriques correspondants sont indiqués dans le tableau ci-dessous :

$t_{i-1} - t_i$	e_i	o_i
0 - 513,96	6	4,8
513,96 - 1027,92	5	4,8
1027,92 - 1541,88	3	4,8
1541,88 - 2055,84	5	4,8
2055,84 - 2569,80	2	4,8
2569,80 - 3083,76	4	4,8
3083,76 - 3597,72	5	4,8
3597,72 - 4111,68	3	4,8
4111,68 - 4625,64	9	4,8
4625,64 - 5139,60	6	4,8
	48	48

Le χ^2 correspondant à 9 degrés de liberté, vaut 7,42.

or $\Pr\{\chi^2 > 7,42\} \# 0,40$.

donc la loi des $T(\tau_i)$ est bien uniforme.

4/ Test 4.

Il consiste à tester si le temps d'apparition de la première avarie est anormalement court.

$$T(\tau_1) = 58,8$$

$$T(\tau_{49} - \tau_1) = 5080,8$$

$$\frac{(r-1)\,T(\tau_1)}{T(\tau_{49} - \tau_1)} = \frac{48 \times 58,8}{5080,8} = F$$

$$\alpha = 0,05 \qquad \Pr\left\{F(2,98) \leqq \frac{1}{3,1}\right\} = 0,05$$

or :

$$T(\tau_1) = 58,8 > \frac{1}{3,1}\,\frac{5080,8}{48} = 34,14$$

donc τ_1 n'est pas anormalement petit.

5/ Test 5.

Il consiste à tester si le temps d'apparition de la première avarie est anormalement long.

$$T(\tau_1) = 58,8 < 3,1\,\frac{5080,8}{48} = 328,11$$

donc τ_1 n'est pas anormalement long.

6/ Test 6.

Il consiste à tester si la durée de vie moyenne calculée sur l'intervalle $(0, \tau_k)$ n'est pas significativement différente de celle calculée sur (τ_k, τ_r).

Prenons :

$$k = 20$$

$$T(\tau_{20}) = 2098,0$$

$$T(\tau_{49} - \tau_{20}) = 3041,6$$

$$F(40,\ 58) = \frac{29 \times 2098,0}{20 \times 3041,6} = 1,000$$

L'écart n'est manifestement pas significatif.

7/ Test 7.

$$kr = 49$$

On forme 7 groupes de 7 valeurs.

Les sommes des durées de vie pour chacun des groupes sont les suivantes :

$T(\tau_7) = 547,4$

$T(\tau_{14} - \tau_7) = 1445,3 - 547,4 = 897,9$

$T(\tau_{21} - \tau_{14}) = 2303,9 - 1445,3 = 858,6$

$T(\tau_{28} - \tau_{21}) = 3239,1 - 2303,9 = 935,2$

$T(\tau_{35} - \tau_{28}) = 4171,8 - 3239,1 = 932,7$

$T(\tau_{42} - \tau_{35}) = 4559,2 - 4171,8 = 387,4$

$T(\tau_{49} - \tau_{42}) = 5139,6 - 4559,2 = 580,4$

Le test d'homogénéité de Bartlett introduit la quantité :

$$\frac{2 \times 49 \left\{ \text{Log}\ \frac{5139,6}{7} - \frac{1}{7} [\text{Log}\ 547,4 + \text{Log}\ 897,9 + \text{Log}\ 858,6 + \text{Log}\ 935,2 + \text{Log}\ 387,4 + \text{Log}\ 580,4] \right\}}{1 + \frac{8}{6 \times 49}}$$

qui se comporte comme un χ^2 à 6 degrés de liberté.

Or cette quantité vaut 4,48 et $\Pr\{\chi^2 > 4,48\} \neq 0,60$ donc on peut conclure à la constance de ϑ entre les différents groupes.

8/ Test 8.

$$r = 1 \qquad \text{donc} \qquad k = 49$$

Les sommes de durées de vie entre avaries successives sont les suivantes :

58,8	67,2	108,5	271,6	67,2	14,0	111,3
48,0	57,4	306,0	32,4	114,0	13,0	7,8
126,9	4,0	16,5	83,2	83,8	139,2	126,0
4,6	339,3	28,8	175,0	88,2	28,6	94,4
81.0	15.2	120,9	271,2	178,5	118,0	153,0
8.8	29 6	72,0	64,4	56,0	59,4	75,2
219,3	385,2	205,9	37,4	45,0	15,2	12,7

ν (nombre de degrés de liberté du χ^2) étant égal à 48 on considère la variable : $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ qui suit approximativement une loi normale centrée réduite.

Elle est égale à - 0,39 et $\Pr\{t > -0,39\} = 0,6517$ donc on conclut également à la constance de ϑ.

9/ Test 9.

$$F_{max} = \frac{\max T(\tau_{kr})}{\min T(\tau_{kr})}$$

D'après les valeurs des $T(\tau_{kr})$ prises pour établir le test 7,

$$F_{max} = \frac{935,2}{387,4} = 2,414$$

On entre dans les tables établies par Hartley avec $k = 7$ et $\nu = 14$.

Le F_{max} correspondant à un seuil α de 5 % vaut 4,95, donc on conclut ici aussi à la constance de ϑ.

10/ Test 11.

La figure 2 montre les fluctuations de $F_n[T(\tau)]$ autour de $F[T(\tau)]$. Elles se trouvent entièrement à l'intérieur de la bande (L_1, L_2) définie par :

$$\Pr\{\max |F_n[T(\tau)] - F[T(\tau)]| \leqslant 0,1941\} = 0,95$$

(valeur lue dans la table de la statistique de Kolmogorov [V]).

On accepte donc l'hypothèse faite sous forme analytique :

Le calcul direct de la formule confirme, évidemment, cette conclusion :

La déviation maximum se produit pour l'abscisse $T(\tau_{33})$ et comme $T(\tau) = 4070,8$ elle est égale à :

$$\left| \frac{32}{48} - \frac{4070,8}{5139,6} \right| = 0,126$$

Or, on a vu que :

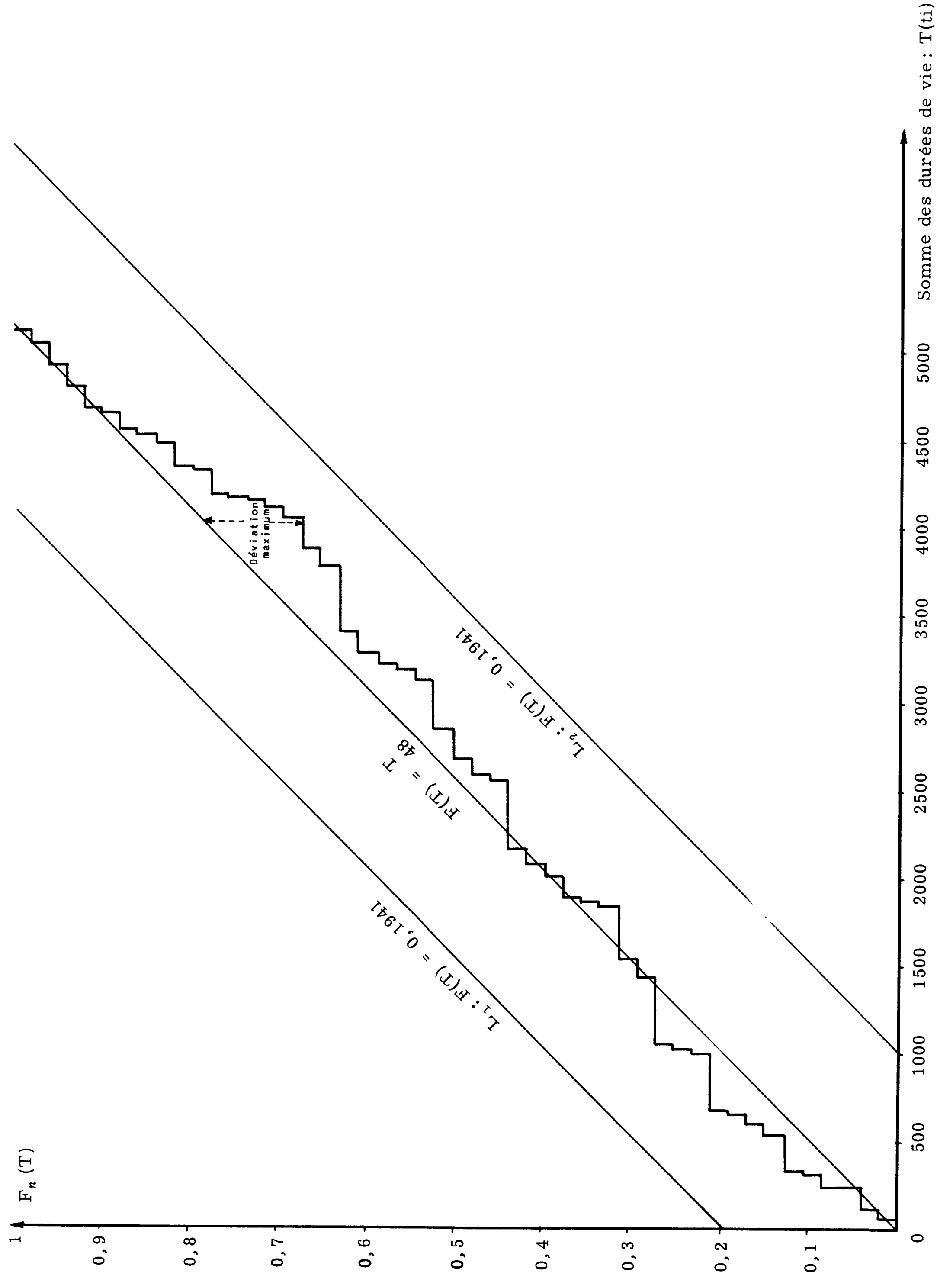
$$\Pr\{\max |F_n[T(\tau)] - F[T(\tau)]| \leqslant 0,1941\} = 0,95$$

Donc on accepte l'hypothèse faite que la distribution des avaries est de Poisson, de paramètre λ constant.

11/ Test 12.

On forme les rapports $\frac{n_k}{N - n_1 - \ldots - n_{k-1}}$ [où n_k désigne le nombre d'avaries observées dans l'intervalle $((k-1)T, kT)$] qui doivent fluctuer autour de la valeur ϑ.

On a pris $T = 40$ et obtenu les résultats suivants :



F_n (T)
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
Somme des durées de vie : T(ti)
L₁ : F(T) = 0,1941
F(T) = T
48
L₂ : F(T) = 0,1941
Déviation maximum

Fig. 2

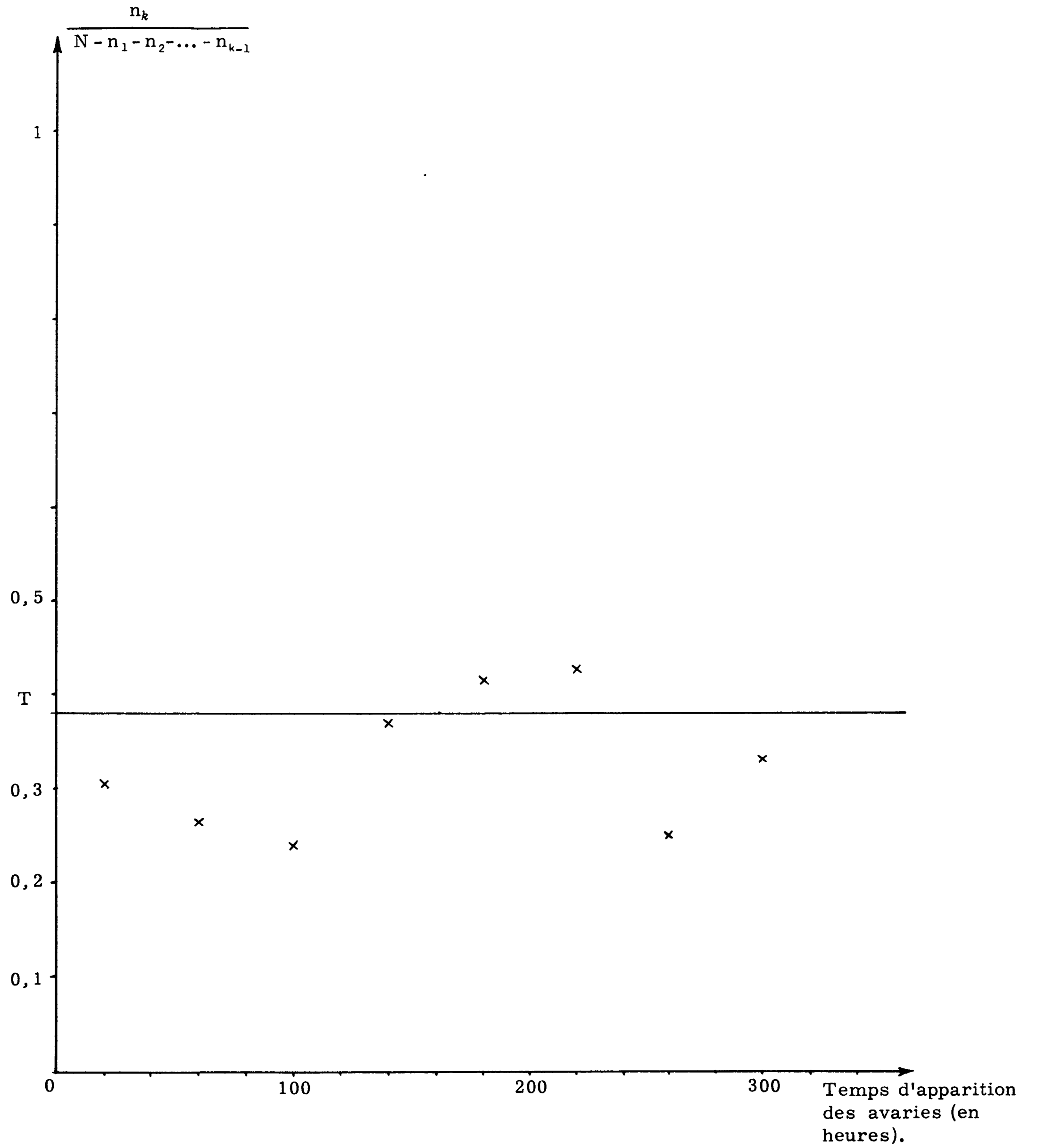

$\frac{n_k}{N - n_1 - n_2 - \ldots - n_{k-1}}$
1
0,5
T
0,3
0,2
0,1
0
100
200
300
Temps d'apparition des avaries (en heures).

Fig. 3

$\frac{2i-1}{2}T$	$N-n_1-\ldots--n_{i-1}$
20	0,306
60	0,264
100	0,240
140	0,368
180	0,416
220	0,428
260	0,250
300	0,333
340	1

et $\frac{T}{\hat{\vartheta}} = 0,381$

La figure 3 montre les positions respectives des rapports par rapport à la valeur théorique.

On peut calculer le χ^2. Il vaut 0,1634 avec 7 degrés de liberté (on a éliminé la dernière classe à laquelle correspond toujours un rapport égal à l'unité).

La probabilité conditionnelle d'avarie d'un matériel dans l'intervalle $(t, t + \Delta t)$ est donc bien indépendante de t.

Nota : Les tests 10 et 13 n'ont pas été consignés faute de pouvoir trouver les tables donnant les valeurs nécessaires.

BIBLIOGRAPHIE

[I] B. EPSTEIN - "Exponential distribution and its role in life testing" Industrial Quality Control, Déc. 1958.

[II] B. EPSTEIN - "Tests for the validity of the assumption that the underlying distribution of life is exponential" - Technometrics, vol.2 - n° 1 - Février 1960.

[III] B. EPSTEIN - "Tests for the validity of the assumption that the underlying distribution of life is exponential" - Technometrics, vol.2 - n° 2 - Mai 1960.

[IV] R.F. DRENICK - "Failure law of complex equipment" - J. Soc. Industr. Appl. Math. Vol.8, n° 4 - Déc. 1960.

[V] Z.W. BIRNBAUM - "Numerical tabulation of the distribution of Kolmogorov's Statistic for finite sample size" - Journal of the American Statistical Association, Vol.47 - N° 259 - Sept.1952.

[VI] R. DESCAMPS - "Durée de vie et défaillances des matériels" - Bulletin de l'A.F.C.I.Q., Décembre 1961.

[VII] B. EPSTEIN - "Recents developments in life testing" - Bulletin Inst. Internat. de Stat. 33ème session, Paris 1961 (avec une importante bibliographie).

[VIII] G. SCHWARZ - "Sequential life testing" - Bulletin Inst. Internat. de 33ème session, Paris 1961.

[IX] B. SHERMAN - "A random variable related to the spacing of sample values" - Ann. Math. Stat. 21 - N° 3 - 1950.

[X] D.J. BARTHOLOMEW - "Note on the use of Sherman's statistic as a test of randomness" - Biometrika 41, 556-558, 1954.

[XI] D.J. BARTHOLOMEW - "Testing for departure from the exponential distribution" - Biometrika 44, 253-256, 1959.